

Conceptual models for markup structured documents : schema languages and formatted-processable documents

著者	村田 真
内容記述	Thesis (Ph. D. in Engineering)--University of Tsukuba, (B), no. 2222, 2006.7.25 Includes bibliographical references
発行年	2006
URL	http://hdl.handle.net/2241/18606

氏 名（本籍）

むら

た

まこと

村 田 真（神奈川県）

学位の種類

博士（工学）

学位記番号

博 乙 第 2222 号

学位授与年月日

平成 18 年 7 月 25 日

学位授与の要件

学位規則第 4 条第 2 項該当

審査研究科

システム情報工学研究科

学位論文題目

Conceptual Models for Markup Structured Documents : Schema Languages and Formatted-Processable Documents

（マークアップ構造化文書のための概念モデル：スキーマ言語と書式付き処理可能文書）

主 査

筑波大学教授

理学博士

北 川 博 之

副 査

筑波大学教授

理学博士

大 保 信 夫

副 査

筑波大学教授

Ph. D.

田 中 二 郎

副 査

筑波大学講師

博士（理学）

南 出 靖 彦

副 査

京都大学教授

工学博士

吉 川 正 俊

論 文 の 内 容 の 要 旨

本論文は, SGML または XML 文書などのマークアップ構造化文書のための概念モデルを扱っている。マークアップ構造化文書とは, 木構造などの構造を備えた構造化文書を, マークアップを含んだテキストによって表現したものである。

前半部では, 正規木文法に基づいて, XML スキーマ言語のための形式的なフレームワークを示している。本フレームワークによって, スキーマ言語の意味を形式的に定義すること, スキーマ言語の記述能力を比較すること, 検証アルゴリズムを論じることが可能になる。主な研究内容は次のとおりである。第一は, 4つのクラスの木言語（「局所的」, 「単一型」, 「抑制された競合」, 「正規」）を比較検討するためのフレームワークの定式化である。第二は, これらのクラスに基づいたスキーマ言語（DTD, W3C XML Schema および RELAX NG）の分類および比較である。第三は, これらのクラスを扱う効率的な文書検証アルゴリズムに関する考察である。第四は, 完全な検証を必要としないデータ型割り当てである。第五は, XML モデルに関連する他の文法概念および高度な検証アルゴリズム（例えば, 二進木への変換による検証, デリバティブに基づく検証）に関する検討である。

後半部では, 文書割付け処理モデルに基づいて, 論理構造と割付け構造の両方を備えた文書を表現するためのマークアップ言語を提案している。同様の目的をもつ SGML の CONCUR は, 論理構造と割付け構造の間でのテキスト順序の逆転を扱えなかった。一方, 提案するマークアップ言語は, ほとんどの WYSIWYG 文書（例えば脚注を含む文書）を表現することができる。後半部の鍵となるアイデアは, 埋込みノードとモールドノードの導入によって, 論理構造と割付け構造の間の関係を解きほぐすことである。埋込みノードは, 論理構造と割付け構造の間の対応関係を表現するノードである。モールドノードは, 一時

的なノードであり、後で割付けノードによって置換えられるノードである。論理構造と割付け構造の間の関係を解きほぐすことにより、論理構造と割付け構造を埋込みマークアップによって容易に表現することができる。

審 査 の 結 果 の 要 旨

本論文は、XML スキーマ言語を比較検討するための形式的なフレームワークと、文書割付け処理モデルに基づく文書の論理構造と割付け構造の表現について考察したものである。前者は、その理論的な意義のみでなく、スキーマ言語（特に、RELAX NG）の標準化に大きな影響を与える等、実際の面でも有用性が認められ高く評価できる。後者は、現在でも完全な解決がなされていない問題に対する先駆的な研究として意義が大きい。よって、本研究はマークアップ構造化文書に関わる情報工学上の貢献が大きいと認められる。今後、本論文で示された理論的基盤に立脚したシステム実装等が進展することを期待する。

よって、著者は博士（工学）の学位を受けるに十分な資格を有するものと認める。